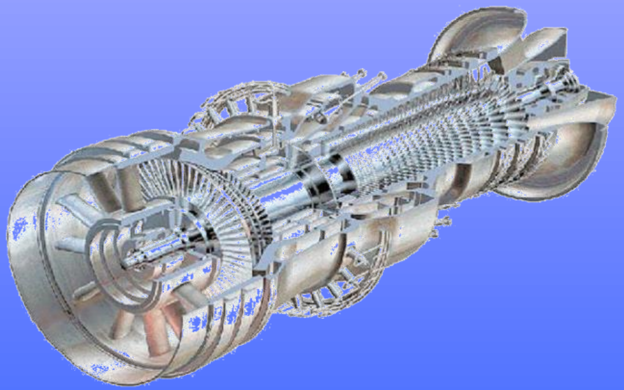
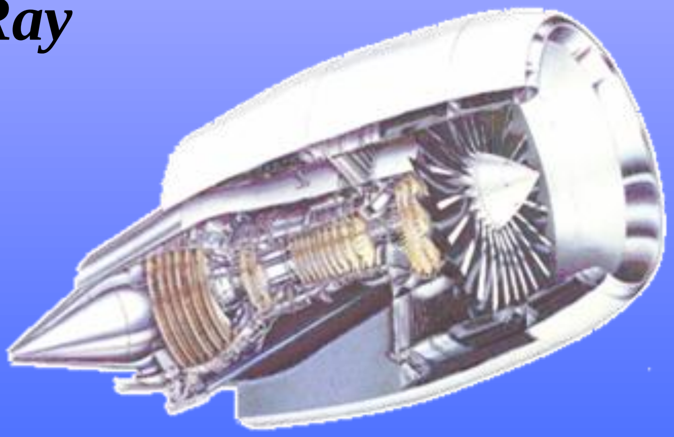
Data-driven Component level Fault Detection in Aircraft Gas Turbine Engines

Soumik Sarkar
Asok Ray



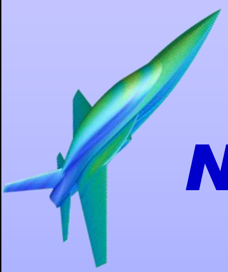


Propulsion Controls and Diagnostic (PCD) Workshop

Cleveland, December 8-10, 2009

2006 NRA TOPIC

IVHM-1.1 *Computationally Efficient High-Fidelity Methods for Propulsion Fault, Degradation, or Damage Detection and Isolation*



NASA Grant and Cooperative Agreement No. NNX07AK49A

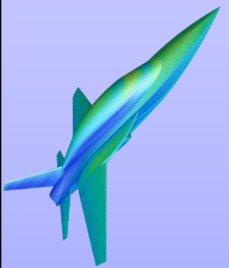
Research Project Title:

Health State Assessment and Failure Prognosis of Integrated Aircraft Propulsion Systems

Technical Monitor (NASA): **Donald Simon <Donald.L.Simon@grc.nasa.gov>**

Principal Investigator (Penn State): **Asok Ray <axr2@psu.edu>**

Specific Contributions of the Current Effort (relevant to *NASA Milestones: 1.2.2.6, 2.3.2.1*)



First year:

➢Fault Detection and Isolation using Symbolic Dynamic Filtering (SDF)

1. S. Gupta, A. Ray, S. Sarkar and M. Yasar, "Fault Detection and Isolation in Aircraft Gas Turbine Engines: Part I – Underlying Concept," ***Proceedings of the I Mech E Part G -- Journal of Aerospace Engineering***, Vol. 222 (G3), No. 3, 2008, pp. 307-318.

2. S. Sarkar, M. Yasar , S. Gupta, A. Ray and K. Mukherjee, "Fault Detection and Isolation in Aircraft Gas Turbine Engines: Part II – Validation on a Simulation Test Bed," ***Proceedings of the I Mech E Part G -- Journal of Aerospace Engineering***, Vol. 222 (G3), No. 3, 2008, pp. 319-330.

➢Comparison of SDF with other Data-driven Engine Diagnostic Tools (e.g. KRA, PCA, UKF, ANN, PF)

3. C. Rao, A. Ray, S. Sarkar and M. Yasar, "Review and Comparative Evaluation of Symbolic Dynamic Filtering for Detection of Anomaly Patterns," ***Signal, Image, and Video Processing,*** Vol. 3, Issue 2 (2009), pp. 101-114.

Second year:

➢Statistical Estimation of Simultaneously occurring Multiple Engine Faults

4. S. Sarkar, C. Rao and A. Ray, "Statistical Estimation of Multiple Faults in Aircraft Gas Turbine Engines," ***Proceedings of the I Mech E Part G: Journal of Aerospace Engineering***, Vol. 223, No. 4, 2009, pp.415-424

➢Generalization of Hilbert Transform: A Theoretical Advancement regarding data preprocessing

5.S. Sarkar, K. Mukherjee and A. Ray, "Generalization of Hilbert Transform for Symbolic Analysis of Noisy Signals," ***Signal Processing***, Vol. 89, Issue 6, 2009, pp. 1245-1251.

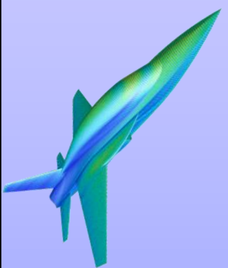
Third year:

➢Symbolic System Identification for On-board Engine Fault Diagnosis

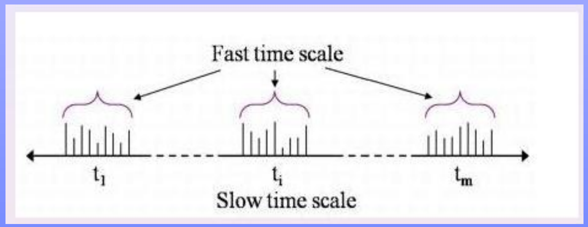
➢Data-driven Diagnostics of Component Faults in presence of Sensor Degradation

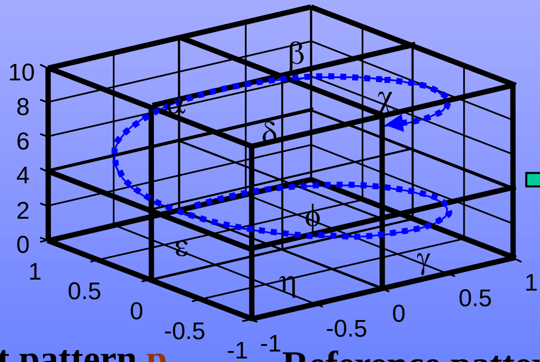
Papers submitted in ACC 2010 and preparation of Journal articles is in progress

Feature Extraction and Pattern Identification from Time Series Data of Macroscopic Observables



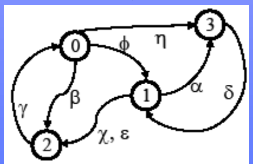
- **Coarse-grained representation of time series data as symbol sequences**
- **Symbol generation and construction of probabilistic finite-state automata (PFSA)**
- **Feature extraction in the form of state probability vectors as patterns**
- **Classification of statistical behavior patterns based on extracted features**
- **Identification of forthcoming departures from the nominal conditions (e.g., major faults and operational disruptions as analog of thermodynamic phase transitions)**





......ϕ χ γ η δ α δ χ......

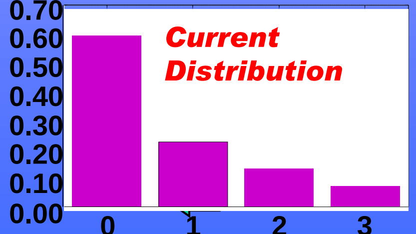
Symbol Sequence



Finite State Machine (Hidden Markov Model)

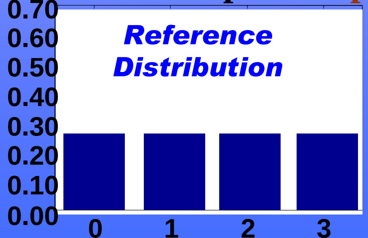
Perron-Frobenius operator (i.e., state transition matrix)

Current pattern p



State Probability Histogram

Reference pattern p^{ref}



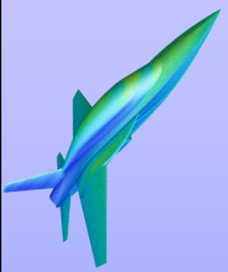
State Probability Histogram

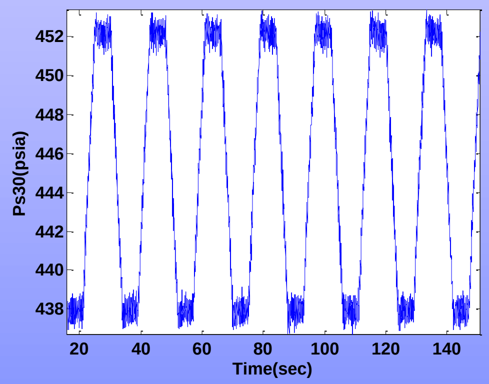
Anomaly measure (i.e., deviation from nominal behavior)

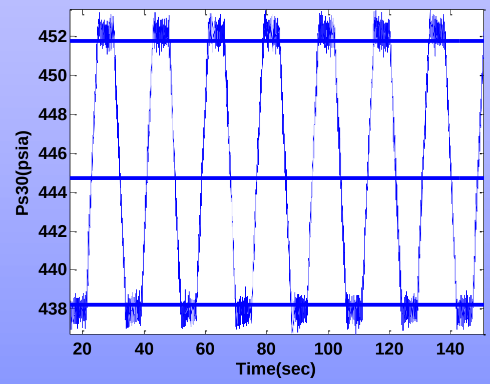
$\mu = d(p, p^{ref})$

Partitioning: A non-linear Feature Extraction Technique

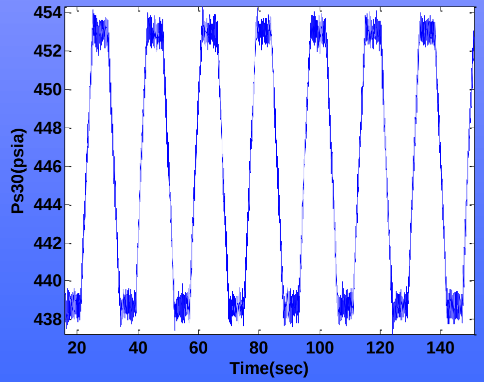
Data Space → **Partitioning Process** → **Pattern Space**

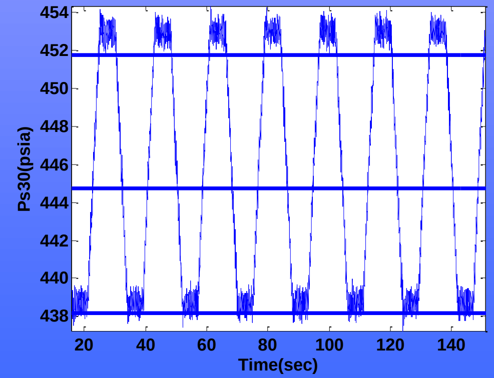






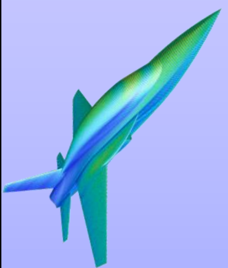
$$\begin{bmatrix} 0.2499 \\ 0.2500 \\ 0.2500 \\ 0.2501 \end{bmatrix}$$



$$\begin{bmatrix} 0.0314 \\ 0.4443 \\ 0.2216 \\ 0.3027 \end{bmatrix}$$

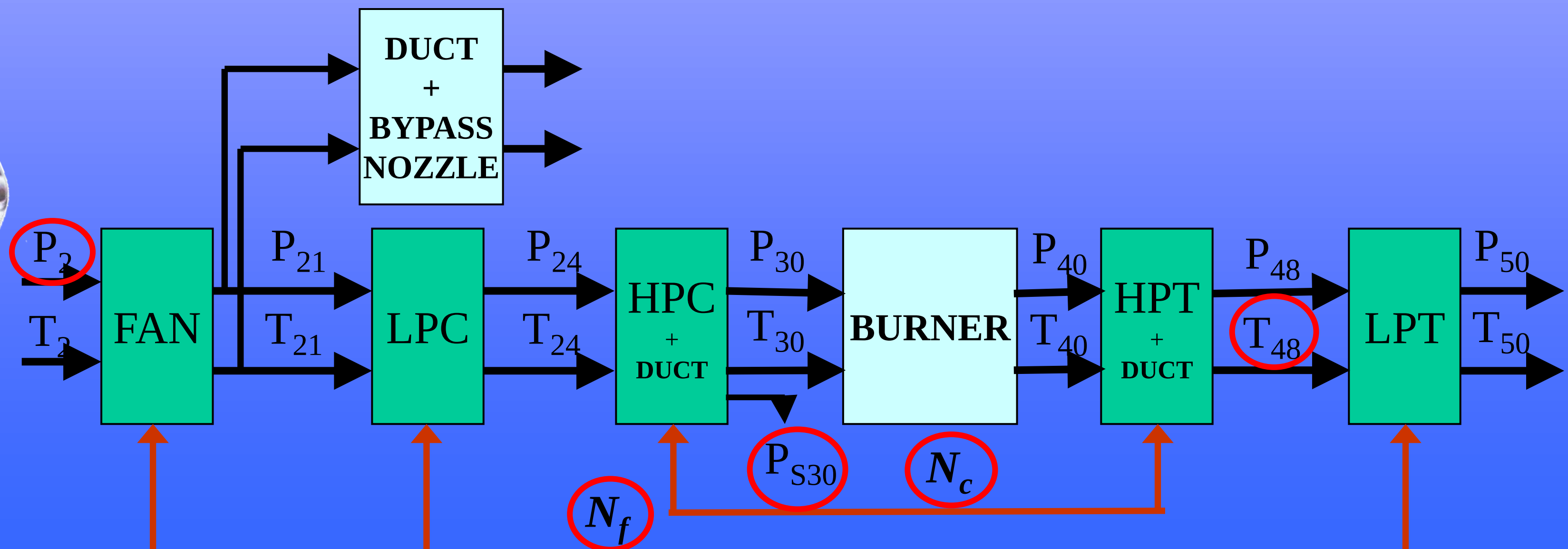
Detection of Component faults in presence of Sensor degradation



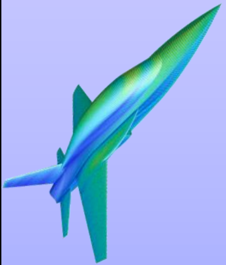

- Sensor degradation can be categorized as
  - ✓ Bias fault
  - ✓ Drifting fault (same as bias upon use of a proper time-scale)
  - ✓ Change in sensor-noise variance

- Traditionally, methods of analytic redundancy are used for sensor failure detection. In the present approach a completely data-driven approach will be pursued to detect component faults in presence of sensor degradation

- Sensors that affect the controller have worse effect on the whole system. In the CMAPSS model, P2, Ps30, T48, N_f, N_C are used by the controller. Degradation (change in sensor-noise variance) in Ps30 and fault in HPC is considered in the present example



Problem Setup for Data-driven Fault Detection

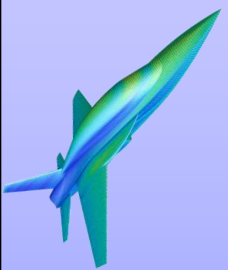


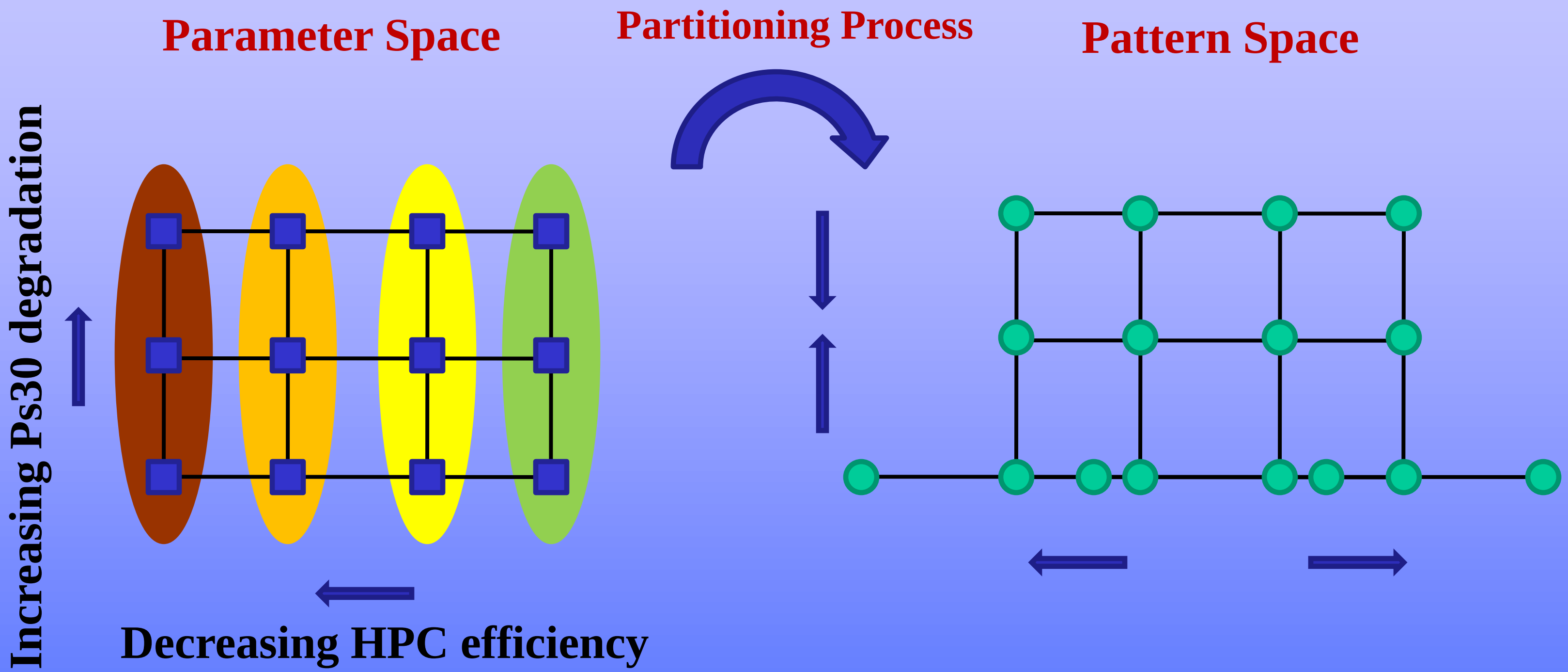
- From the perspective of Detection, we need to identify whether there is any fault in the plant or not. If the data is analyzed blindly, then the detection tool may succumb to a sensor degradation instead of an actual plant fault.

- The idea here is to optimize the detection tool, in the present case, the partitioning process of the SDF feature extraction tool such that, it masks the sensor fault signatures and magnifies the plant fault signatures.

- Provided that the data evolution characteristics due to plant fault and sensor fault are not very similar!

Detection as a Multi-class Classification Problem

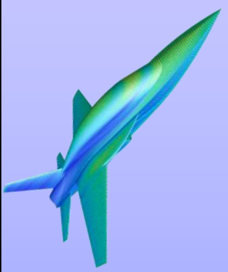


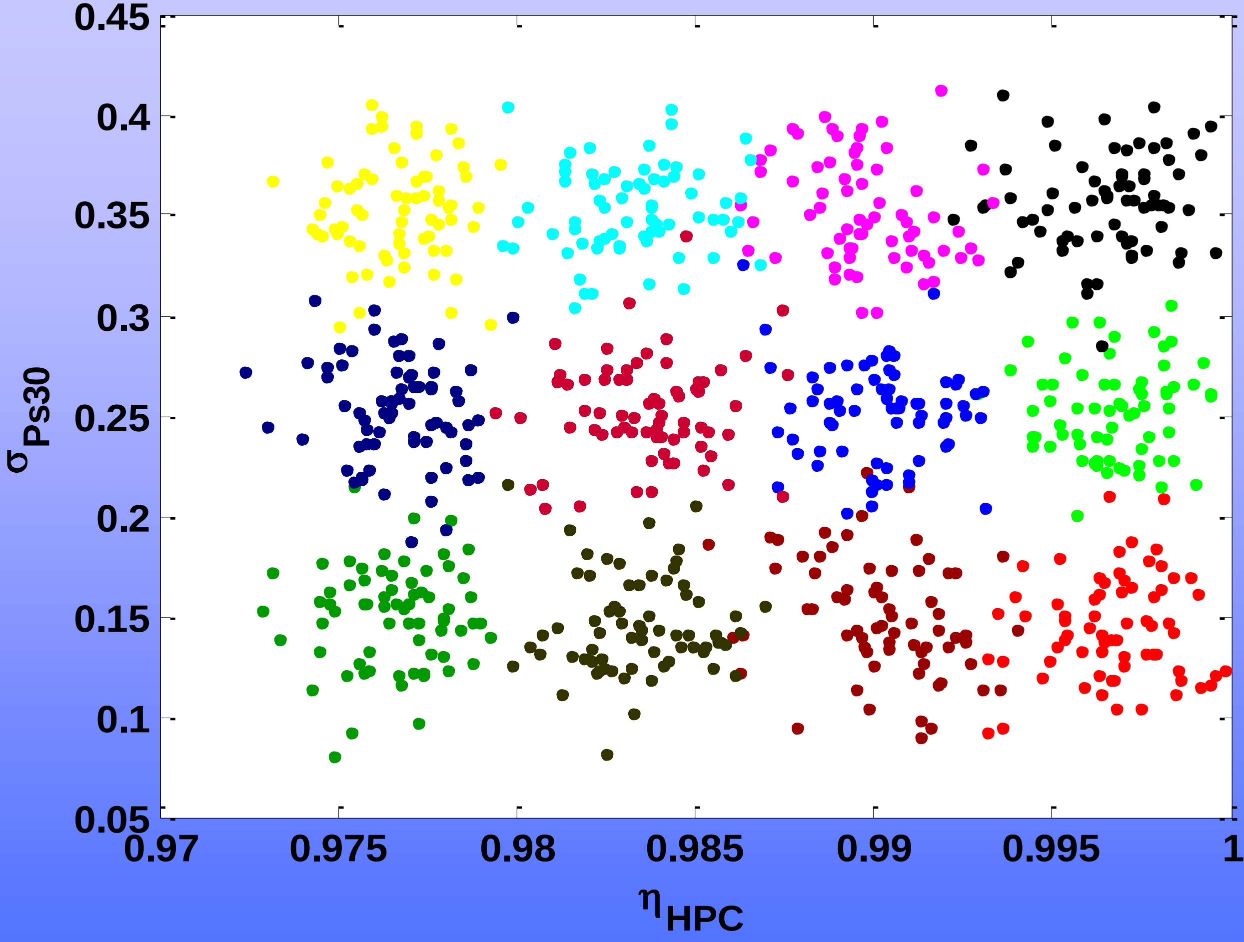


■ **Parameter Value (may be a vector)**

● **Pattern (State probability vector in SDF)**

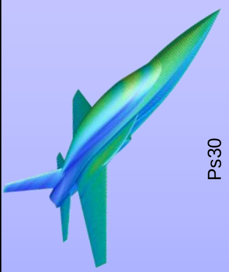
Parameter Space for HPC + Ps30 System

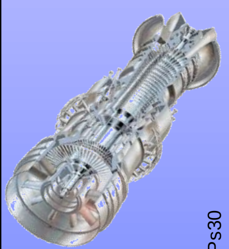


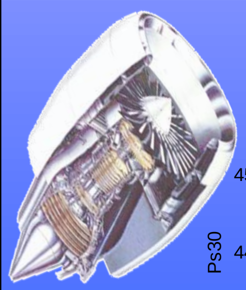


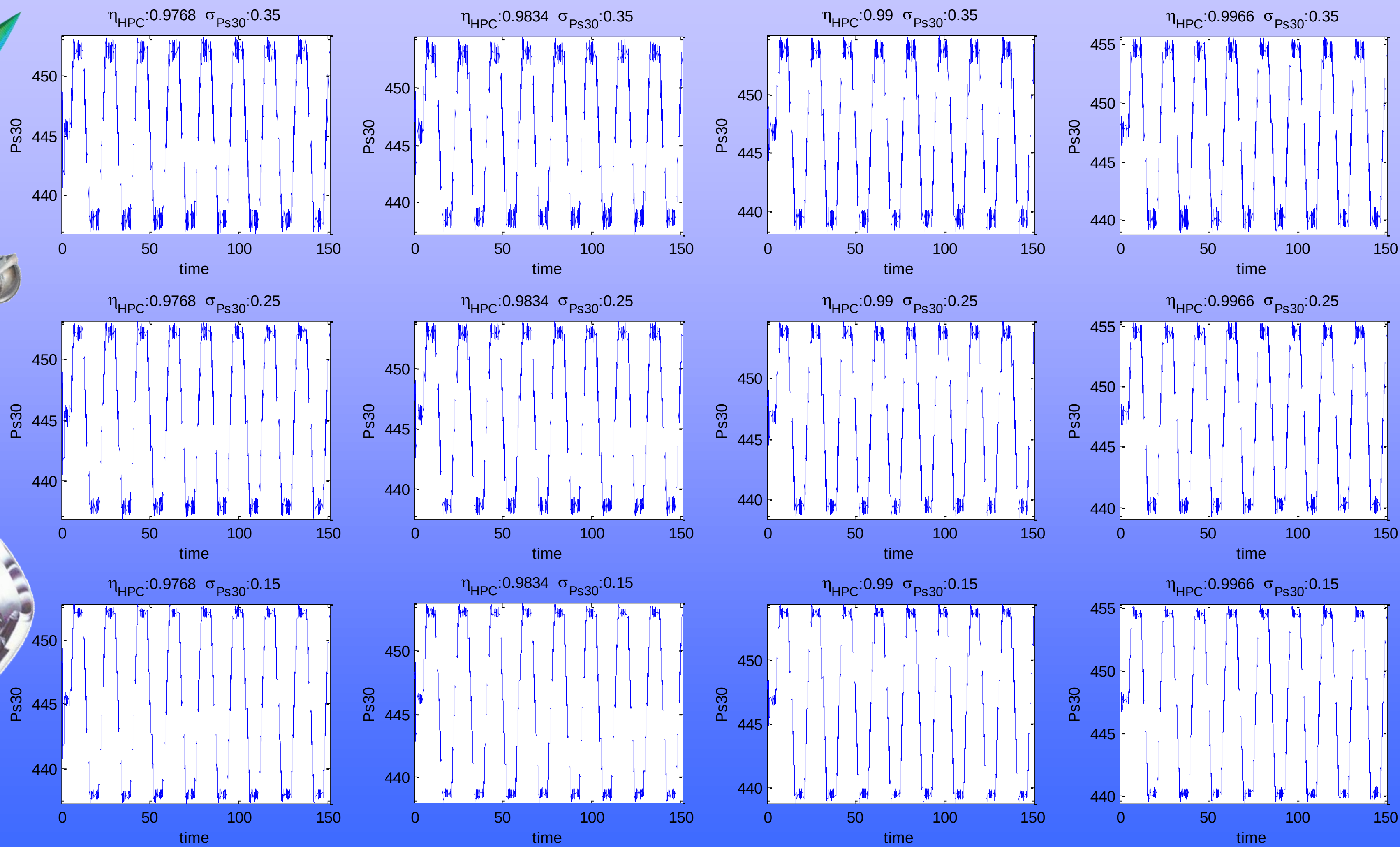
Data is generated using CMAPSS for 12 classes (4 x 3) where parameter values are sampled from Gaussian distributions around the mean parameter values

Representative Data Space

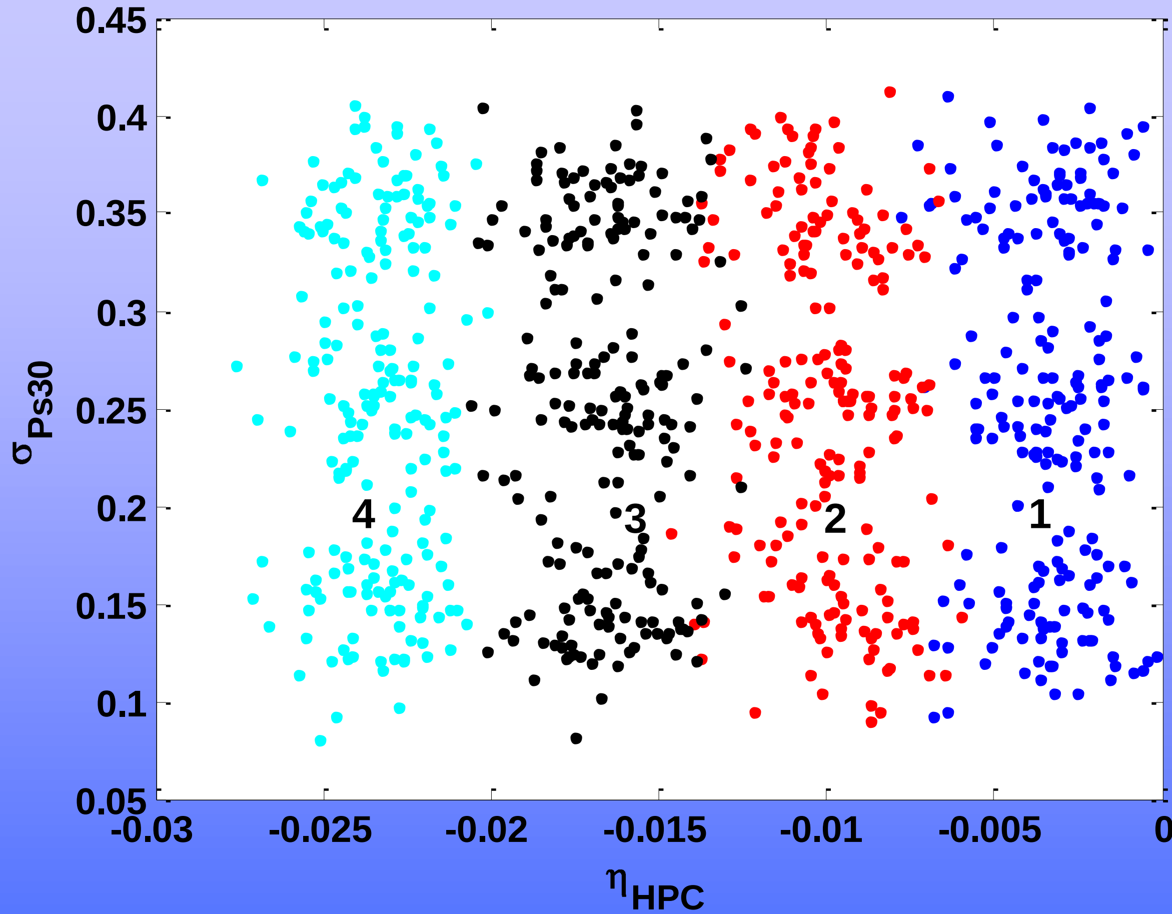






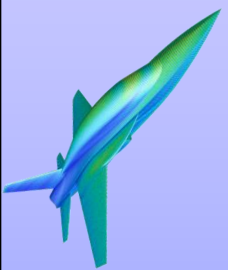


Class Assignment in the Parameter Space

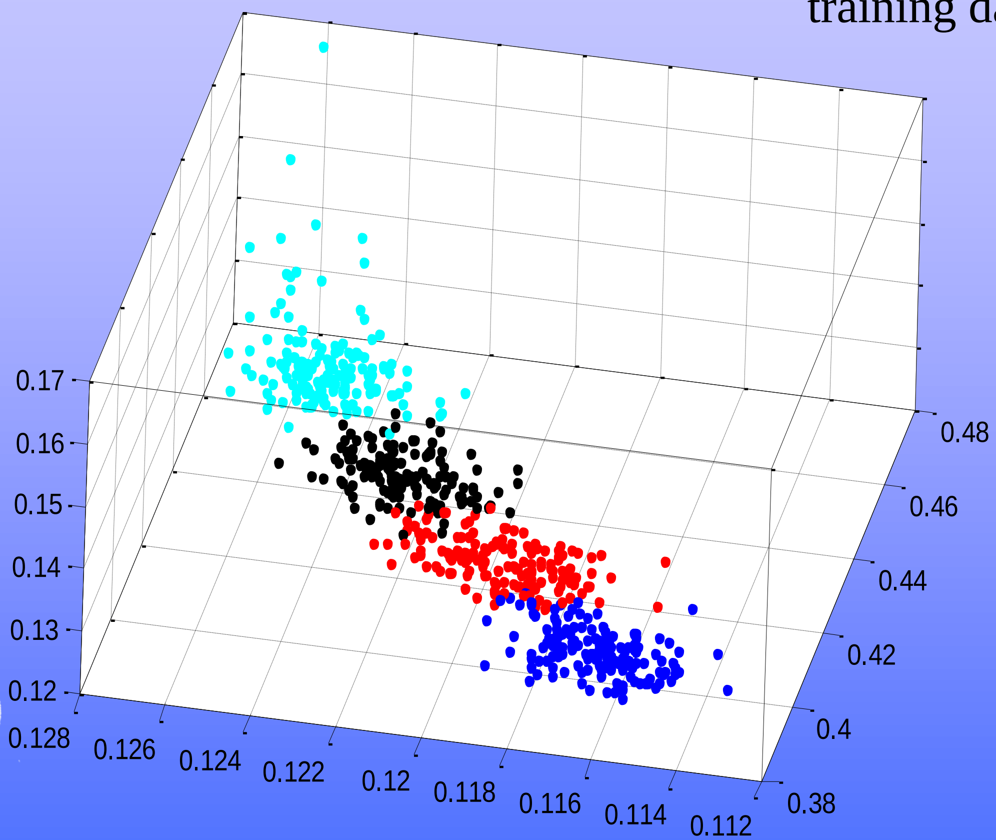


- Magnifying HPC fault signature (assigning different classes across HPC efficiency)
- Masking Ps30 degradation signature (assigning same class across Ps30 degradation)

Uniform Partitioning



Using KNN classifier (K = 5), with training data-set (150 samples per class)



Confusion Matrix

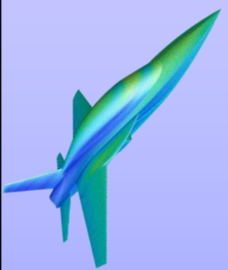
		Estimated			
		C1	C2	C3	C4
Actual	C1	147	3	0	0
	C2	5	140	5	0
	C3	0	5	145	0
	C4	0	0	2	148

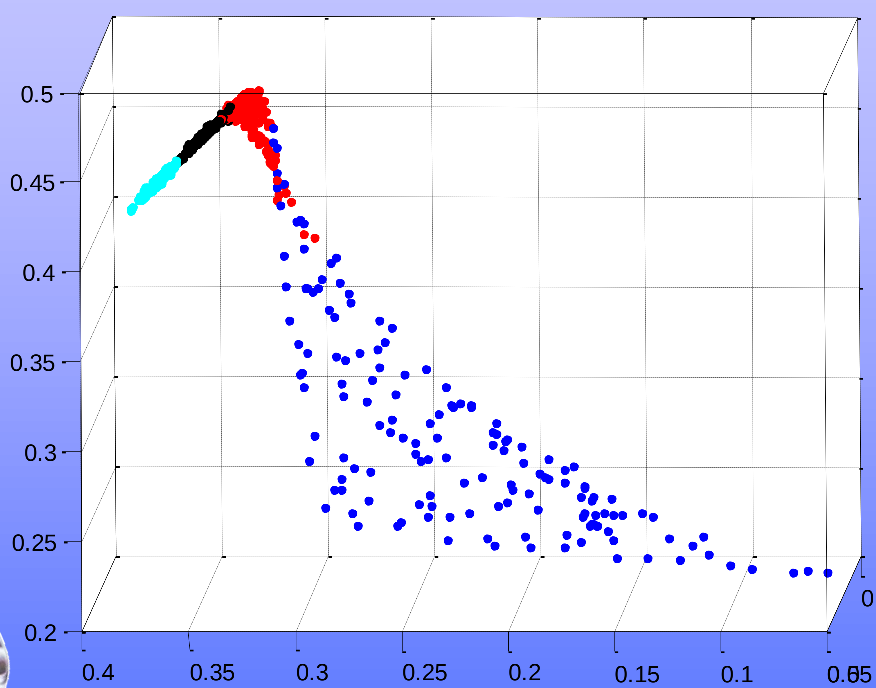
Error %: 3.3333

3-D plot of pattern space

Note the Non-Gaussian nature in the Pattern Space

Maximum Entropy Partitioning





3-D plot of pattern space

Using KNN classifier (K = 5), with training data-set (150 samples per class)

Confusion Matrix

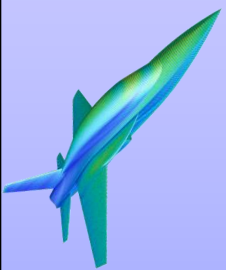
Estimated

Actual	C1	C2	C3	C4
C1	143	7	0	0
C2	3	144	3	0
C3	0	5	145	0
C4	0	0	1	149

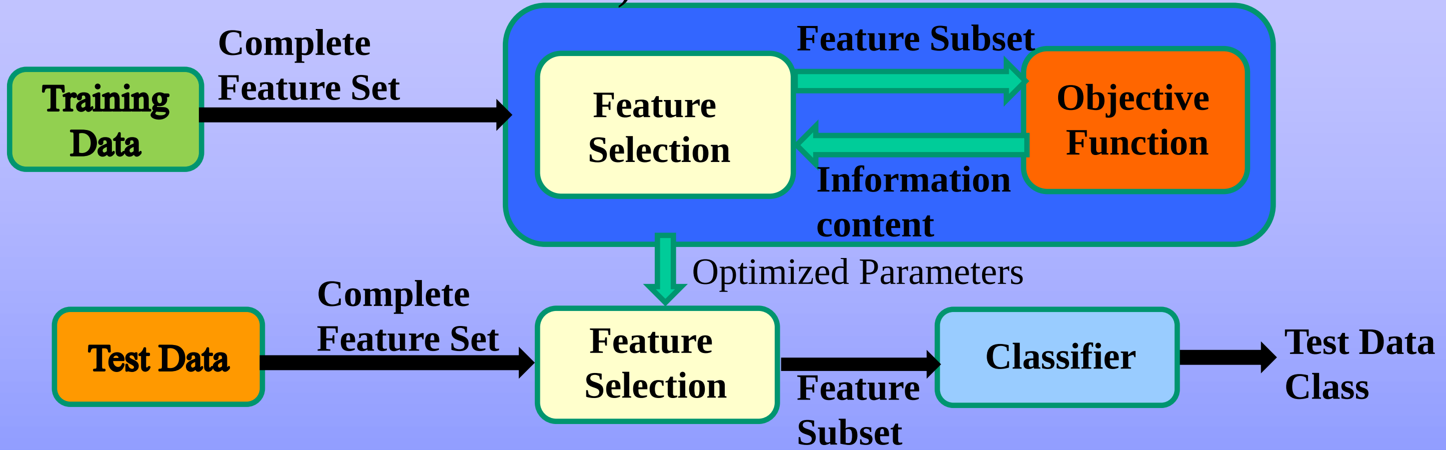
Error %: 3.1667

Note the Non-Gaussian nature in the Pattern Space

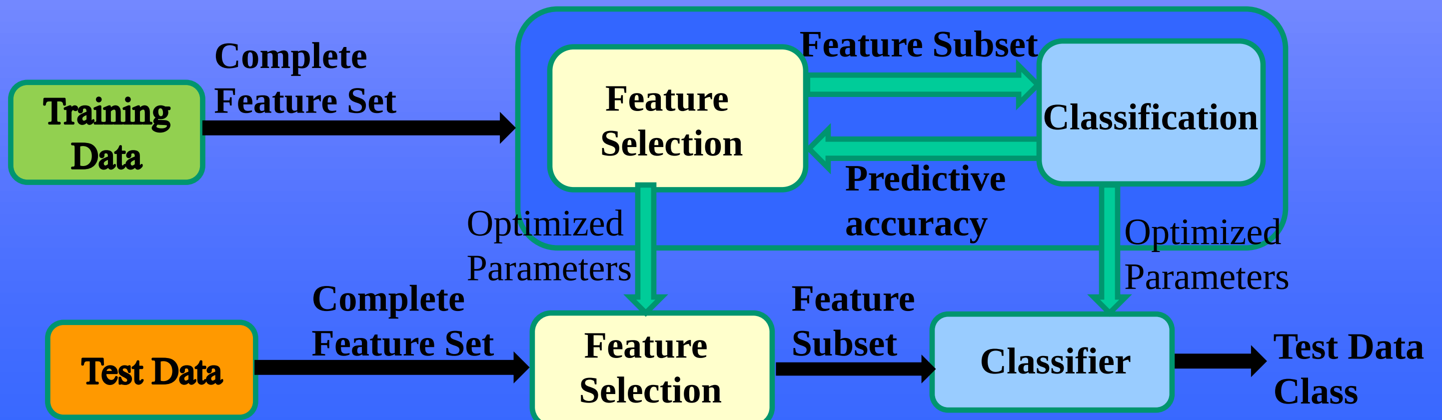
Types of Feature Selection schemes



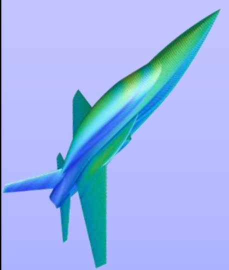
Filters: Uses information content feedback, (e.g. Fisher Variance Ratio, i.e. ratio of Inter-class and Intra-class distance, Statistical Dependence, Information Theoretic measures)

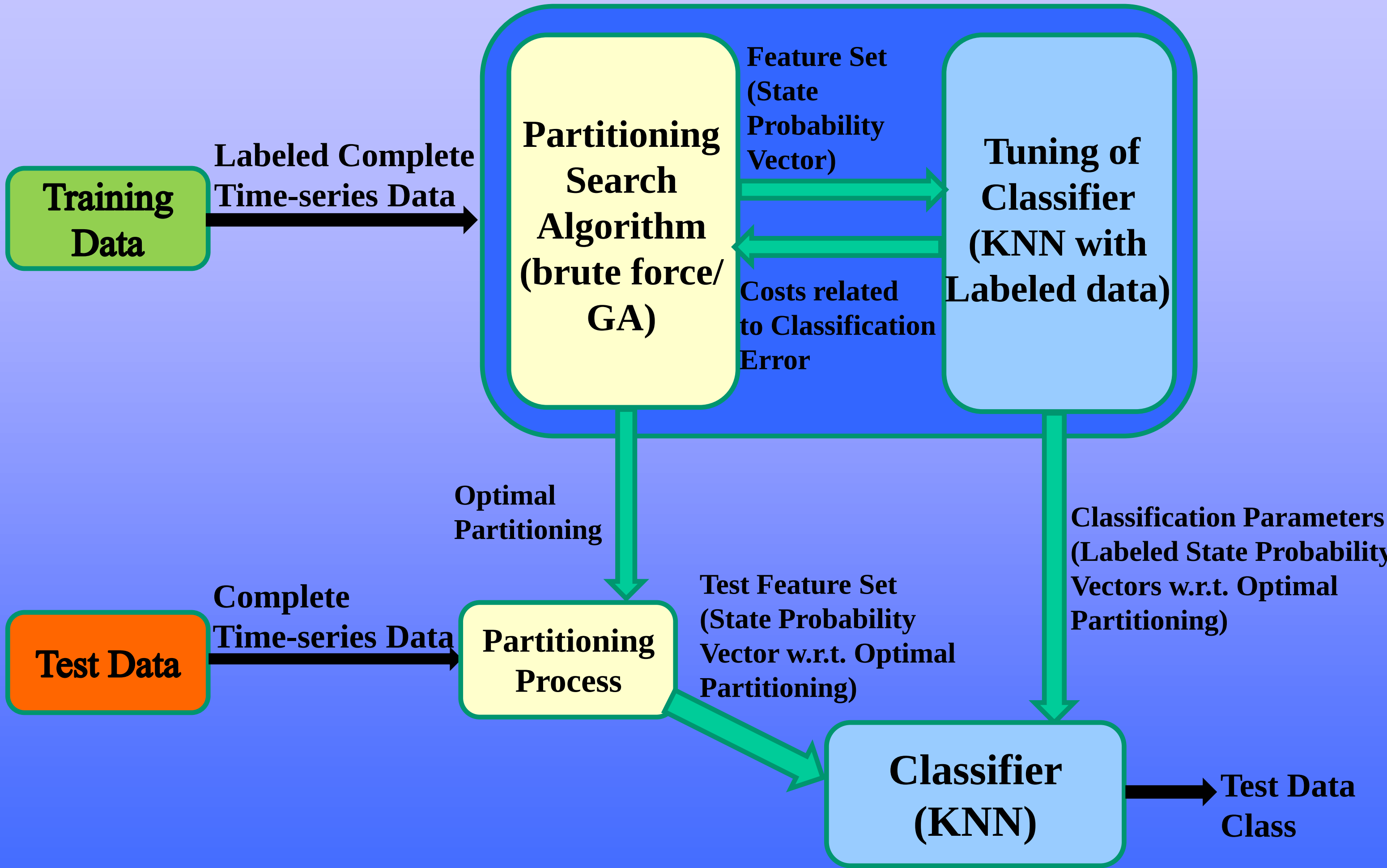


Wrappers: Classifier is included in optimization, objective is to maximize predictive accuracy (e.g. classification rate), using statistical re-sampling or cross-validation

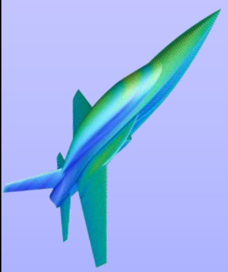


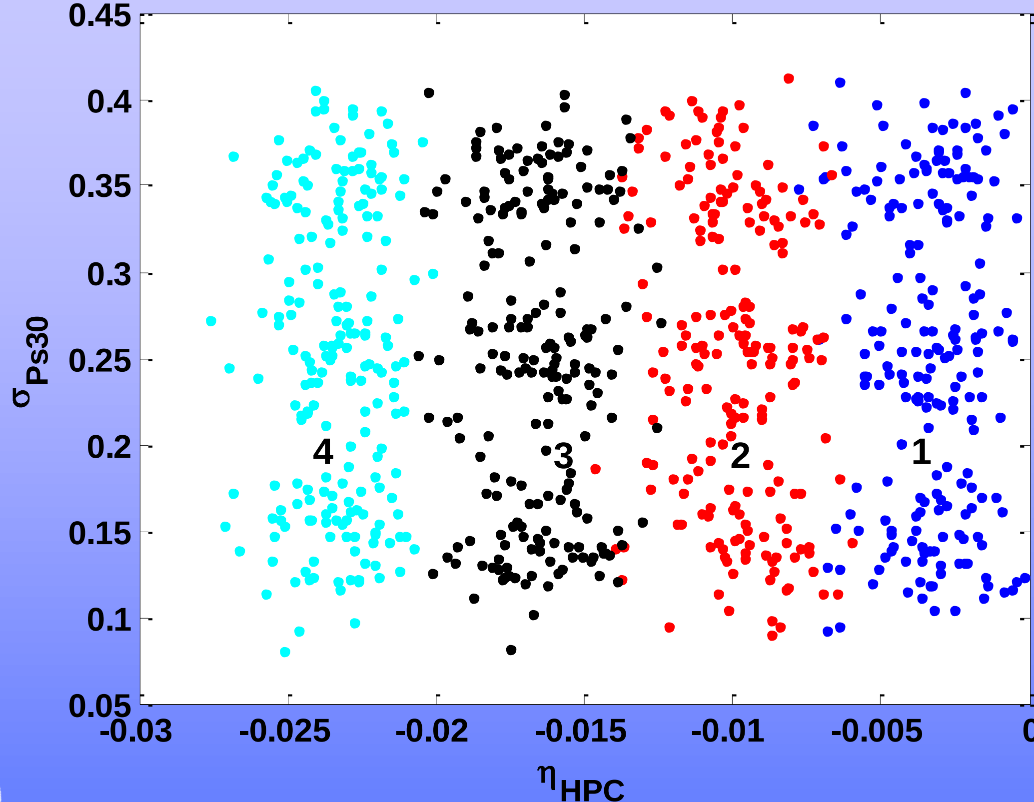
Outline of the Classification Procedure





Multi-objective Cost Function





Confusion Matrix

Estimated(j)

Actual(i)	C1	C2	C3	C4
C1	147	3	0	0
C2	5	140	5	0
C3	0	5	145	0
C4	0	0	2	148

Weighting Matrix

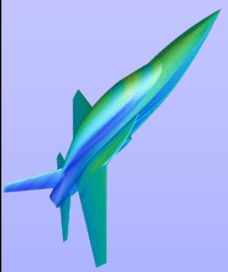
Estimated(j)

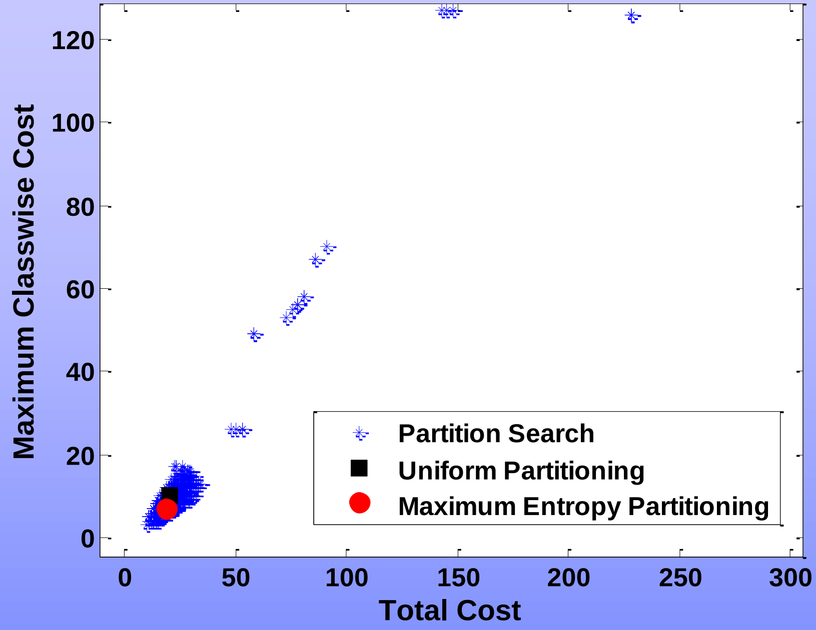
Actual(i)	W1	W2	W3	W4
W1	0	1	2	3
W2	1	0	1	2
W3	2	1	0	1
W4	3	2	1	0

Total Cost $$Cost1 = \sum_i \sum_j W_{ij} C_{ij}$$ (Expected Classification Error)

Maximum Class wise Cost $$Cost2 = \max_i \sum_j W_{ij} C_{ij}$$ (Worst Case Classification Error)

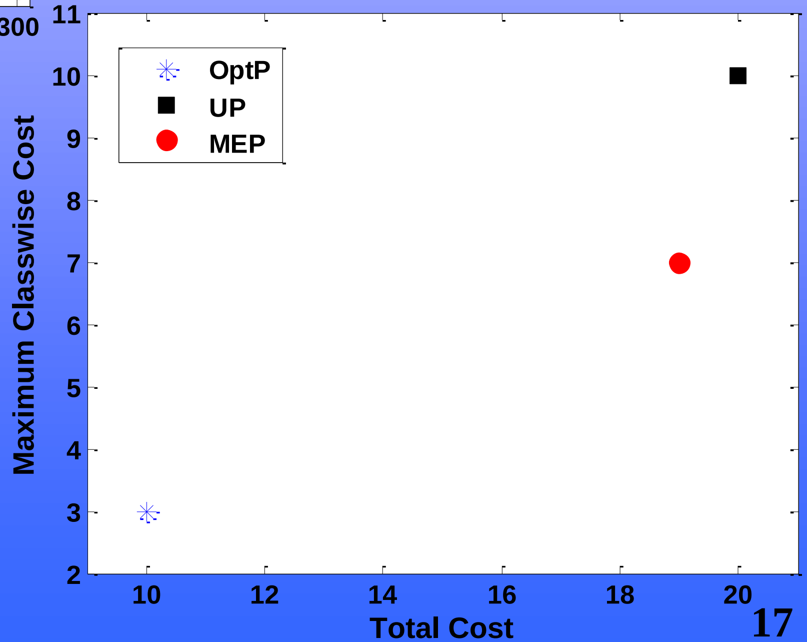
Objective Space and the Pareto Front



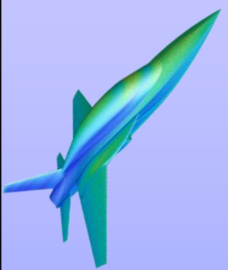



- ➢ Pareto front is generated by identifying the non dominated points
- ➢ Use of Neyman-Pearson criteria for choosing the optimal partitioning

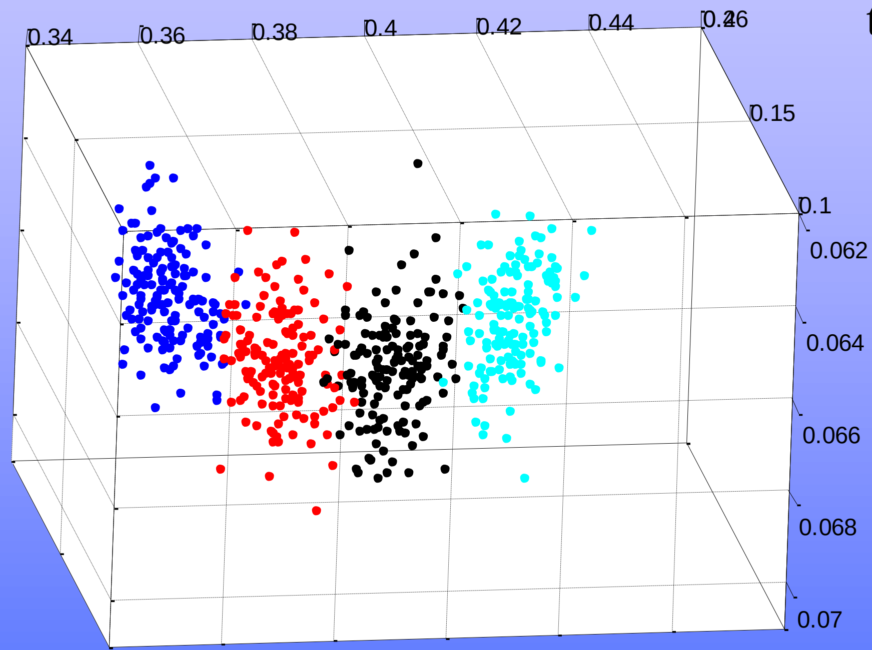
Threshold chosen for Maximum Class wise Cost: 3 (Utopian Point in this case)



Pattern Space and Confusion Matrix for the Optimal Partitioning



Using KNN classifier (K = 5), with training data-set (150 samples per class)



3-D plot of pattern space

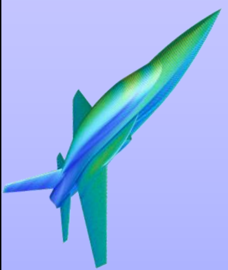
Confusion Matrix

Estimated

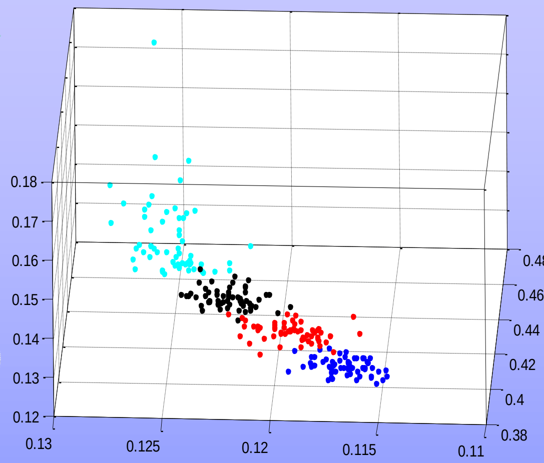
Actual	**C1**	**C2**	**C3**	**C4**
C1	148	2	0	0
C2	0	147	3	0
C3	0	3	147	0
C4	0	0	2	148

Error %: 1.6667

Result on the Test Dataset (60 Samples per class)



Uniform Partitioning

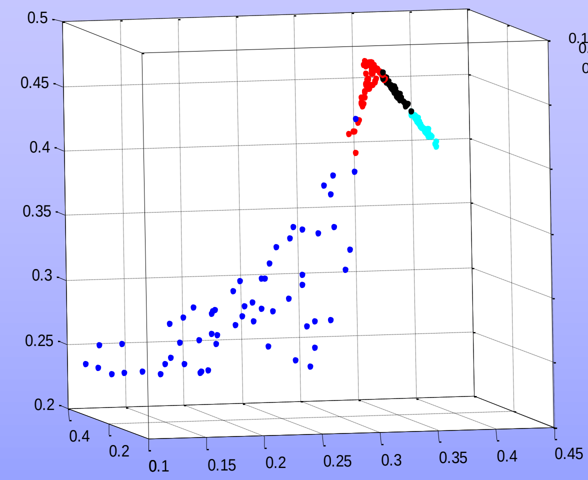


Estimated

Actual	C1	C2	C3	C4
C1	60	0	0	0
C2	3	57	0	0
C3	0	3	56	1
C4	0	0	0	60

Cost1 = 7
Cost2 = 4
Error %: 2.9167

Max. Entropy Partitioning

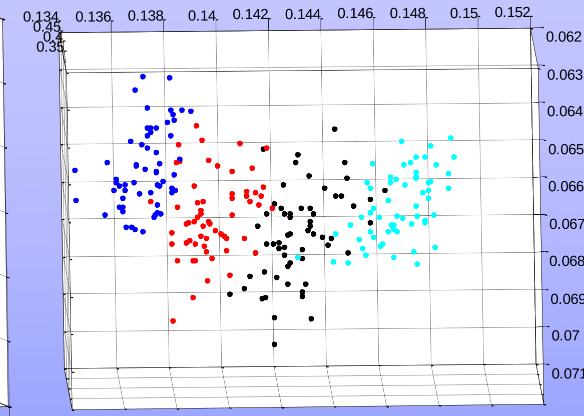


Estimated

Actual	C1	C2	C3	C4
C1	59	1	0	0
C2	2	57	1	0
C3	0	4	55	1
C4	0	0	0	60

Cost1 = 9
Cost2 = 5
Error %: 3.75

Optimal Partitioning

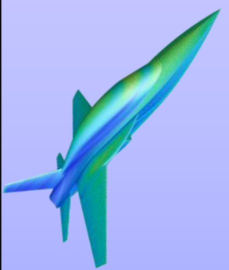


Estimated

Actual	C1	C2	C3	C4
C1	60	0	0	0
C2	1	59	0	0
C3	0	2	58	0
C4	0	0	1	59

Cost1 = 4
Cost2 = 2
Error %: 1.6667

Conclusions

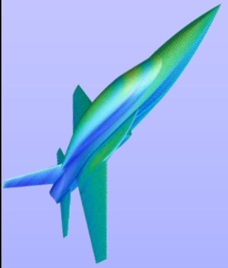


- Data space partitioning as a nonlinear feature extraction technique
- Formulating data driven detection problem of component faults in presence of sensor degradation as a multi-class classification problem
- Construction of a Multi-Objective Cost function for the multi-class classification problem
- Optimizing Partitioning for improving classification rate using the Multi-Objective Cost function

Future Work

- Use of Genetic Algorithm/ Sequential Quadratic Programming etc. for the optimization procedure
- Use of other classifiers (e.g. Support Vector Machines) and comparison of performances

- Validation on other types of sensor degradation
- Investigation of conditions with simultaneous faults in multiple components and sensors
- Developing an information fusion framework to generate composite patterns by fusing atomic patterns from individual sensors



Thank you